Self-adaptive sampling rate data acquisition in JET's correlation reflectometer[a]

G. de Arcas,[1,b] J. M. López,[1] M. Ruiz,[1] E. Barrera,[1] J. Vega,[2] A. Murari,[3] A. Fonseca,[4] and JET-EFDA Contributors[5,c),d)]

[1]*Grupo de Investigación en Instrumentación y Acústica Aplicada, Universidad Politécnica de Madrid, Ctra. Valencia Km 7, 28031 Madrid, Spain*
[2]*Asociación EURATOM/CIEMAT para Fusión, Avda. Complutense 22, 28040 Madrid, Spain*
[3]*Consorzio RFX-Associazione EURATOM ENEA per la Fusione, Corso Stati Uniti 4, Padova 35127, Italy*
[4]*Associação EURATOM/IST, Instituto de Plasmas e Fusão Nuclear, 1049-001 Lisbon, Portugal*
[5]*JET-EFDA, Culham Science Centre, OX14 3DB Abingdon, United Kingdom*

(Presented 14 May 2008; received 12 May 2008; accepted 7 July 2008; published online 31 October 2008)

Data acquisition systems with self-adaptive sampling rate capabilities have been proposed as a solution to reduce the shear amount of data collected in every discharge of present fusion devices. This paper discusses the design of such a system for its use in the KG8B correlation reflectometer at JET. The system, which is based on the ITMS platform, continuously adapts the sample rate during the acquisition depending on the signal bandwidth. Data are acquired continuously at the expected maximum sample rate and transferred to a memory buffer in the host processor. Thereafter the rest of the process is based on software. Data are read from the memory buffer in blocks and for each block an intelligent decimation algorithm is applied. The decimation algorithm determines the signal bandwidth for each block in order to choose the optimum sample rate for that block, and from there the decimation factor to be used. Memory buffers are used to adapt the throughput of the three main software modules (data acquisition, processing, and storage) following a typical producer-consumer architecture. The system optimizes the amount of data collected while maintaining the same information. Design issues are discussed and results of performance evaluation are presented. © *2008 American Institute of Physics.* [DOI: 10.1063/1.2965011]

INTRODUCTION

The shear amount of data to be collected in every discharge of present reactor relevant fusion devices has reached several gigabytes and therefore adequate measures have to be taken to make sure that useless signals are not stored. On the other hand, since fusion devices are still experimental machines, it is difficult to tell *a priori* which part of the signals can be discarded and which one contains the relevant information. To assess the basic quality of the measurements and perform a first real time screening and reduction in the available information, a new adaptive acquisition system based in the ITMS platform developed in cooperation between Universidad Politecnica de Madrid and CIEMAT (Ref. 1) is being tested at JET, which allows adjusting the sampling rate to the bandwidth of the signals. JET's KG8B correlation reflectometry systems are composed of four reflectometer systems,[2,3] that have the fixed frequency channels operating at the frequencies of 76, 85, 92, and 103 GHz, and the corresponding configurable channels in steps of frequency at 76–78, 85–87, 92–96, and 100–106 GHz, respectively. Each channel is equipped with a quadrature phase detector, resulting in 16 reflectometry signals that have to be recorded during the discharge. Before this project these signals were only acquired for 10 s of the plasma discharge at a sampling rate of 2.0 megasamples/s due to memory limitations of the data acquisition systems used. Strong changes in the spectral bandwidth, dependence in the radial and temporal evolution of the CR signals were previously observed,[4] and thus the decision to use this diagnostic to test the architecture here proposed.

SYSTEM DESCRIPTION

The system has been developed with off-the-self commercial hardware components, as shown in Fig. 1. The data acquisition board (NI-PXI 6115) can acquire a maximum of four analog input channels simultaneously with a sample rate up to 10 megasamples/s and 12 bit resolution. The system controller is a 2.0 GHz dual-core embedded controller for PXI (NI-PXI 8105), running Microsoft Windows XP. Both modules are installed in an eight slot chassis (NI-PXI 1042). The external storage system is made of two 10 000 rpm hard drives with serial advanced technology attachment (SATA) interface connected to the system controller through an eSATA Express Card 34 that includes a redundant array of

[a]Contributed paper, published as part of the Proceedings of the 17th Topical Conference on High-Temperature Plasma Diagnostics, Albuquerque, New Mexico, May 2008.
[b]Author to whom correspondence should be addressed. Electronic mail: g.dearcas@upm.es. Tel.: +34913365228.
[c]For a full listing of names and affiliations of the JET-EFDA Contributors, see A. T. Macrander, Rev. Sci. Instrum. **79**, 10F701 (2008).
[d]See the Appendix of M.L.Watkins *et al.*, Fusion Energy 2006 (Proceedings of the 21st International Conference, Chengdu, 2006) IAEA, (2006).

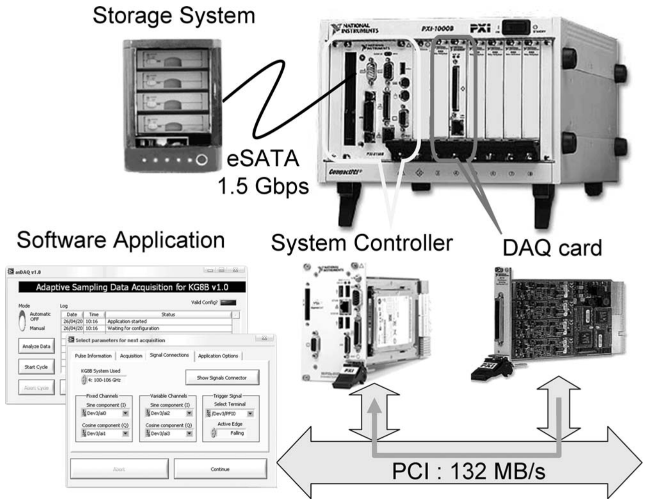

FIG. 1. System components.

inexpensive disks (RAIDs) controller. The hard drives are configured in RAID0 mode providing 150 Gbytes of storage capacity with a maximum sustained write transfer rate of 140 Mbytes/s under Windows XP. The heart of the system is the software application running in the controller, which provides a user interface to set up all acquisition parameters, acquires the data, adjusts the sample rate depending on the signal's bandwidth evolution during the experiment, and stores the results in the external storage system. This application has been developed in LABVIEW Version 8.2.1.

All these hardware and software elements together act as a self-adaptive sample rate (SASAR) data acquisition (DAQ) system,[5] which adjusts the sample rate continuously during the experiment depending on the signal's bandwidth evolution. The process is shown for clarity in Fig. 2. The input signals are acquired at a fixed sample rate that is chosen during setup depending on the expected maximum bandwidth of the signals. Once the acquisition starts the data are read from the DAQ board, split into channels, and placed in a set of memory buffers on the system controller's random access memory. Then, a SASAR algorithm is applied to each data block of every channel. This algorithm estimates the signal's bandwidth for a particular data block for each channel, finds the maximum for all the acquired channels for that block, and decimates the data accordingly to adjust the sample rate to the optimum value for that particular block. Therefore, all input data blocks have the same size, as they have been acquired at a fixed sample rate; but the size of the output data blocks is variable as the sample rate of each one depends on the maximum bandwidth of the signals contained in it. Finally, the output data are placed on a set of memory buffers for storage to disk.

The most critical part of the process is bandwidth estimation, as it is the most time consuming task and therefore limits the system's maximum processing capacity,[6] and consequently the maximum sample rate that can be achieved. It also determines the overall system behavior, as the amount of information included in the output data depends on the qualitative behavior of this algorithm. It must be noted that for each block the system will discard all the information above the estimated bandwidth. Therefore, special attention

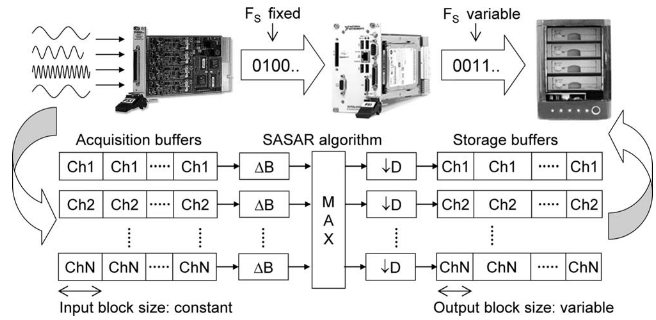

FIG. 2. Software implementation scheme of a SASAR data acquisition system.

must be taken when choosing this algorithm, both from a qualitative point of view and regarding its implementation.

From the point of view of the system's behavior, the first problem encountered when designing this algorithm is to define when the system must consider that a frequency component is carrying information, and when it must consider it as noise. To clarify this situation Fig. 3 shows various examples of the typical shapes of the power spectra found in a KG8B acquisition. In cases (a)–(c) it might be easy to decide the frequency above which the signal is not providing valuable information. However, case (d) might not be so clear without additional information because of the spikes that appear at the fifth to the eighth harmonics of 100 kHz. These spikes can be noise produced in the electronics, in which case they do not provide information concerning the study of the plasma properties, and thus must not be taken into account when estimating the bandwidth or might be produced because of some change in the plasma properties, in which case they are valuable information, and therefore must be included in the estimation. One algorithm has been developed for each interpretation.

The first algorithm, named noise corner frequency estimator (NOCOFE), computes the signal bandwidth as the frequency point where the level of the power spectra meets the noise level. This is done in the following steps: first the single-sided autopower spectrum is obtained; then a low pass filter is applied to smooth it; the noise level is obtained as the mode of the last 20% of the power spectrum; and finally, starting from the power spectrum peak, the frequency (noise

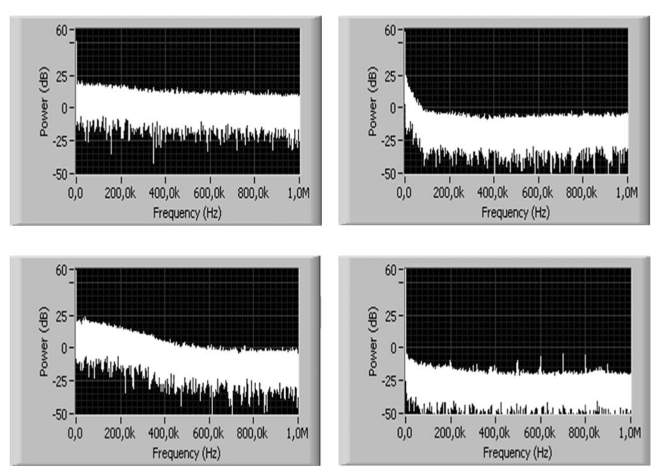

FIG. 3. Typical shapes of the power spectra found in a KG8B acquisition.

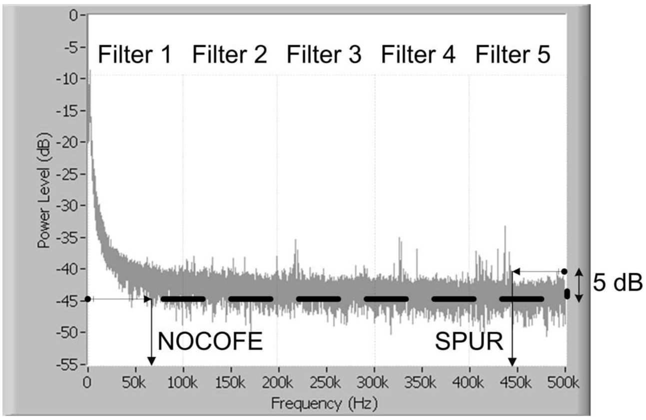

FIG. 4. Example of bandwidth estimation for the three algorithms developed.

corner) at which the power meets the noise level is found. The algorithm applies an excess factor of 10% to the estimated bandwidth and limits its value between 10% and 50% of the sample rate used during acquisition. The lower limit is imposed in order to guarantee that a minimum amount of data is generated for every block. This would be valuable to detect a fault in the system behavior (if the power spectrum of the generated data has been decimated and does not finish in a small flat region the system might have discarded relevant information).

The second algorithm, named SPUR, is a slight modification of the abovementioned that takes into account the possible spurious that might appear in the power spectrum when estimating the bandwidth. This is done by changing only the way of "searching:" in this case the search starts from the end of the power spectrum and finishes on the first point with a level higher than the noise level plus a predefined margin (5 dB). Figure 4 shows an example of how the NOCOFE and SPUR algorithms would estimate the bandwidth of a data block using the smoothed power spectrum: the thick dashed line represents the calculated noise level; the search mechanism of each algorithm is represented with a line that has a circle at the starting point and an arrow at the finish point. This figure shows how different the results from both algorithms can be, and thus the importance of choosing the correct one depending on which type of information is considered relevant. Two more algorithms have been developed based on filter banks due to their more efficient implementation. FILTERB uses uniform frequency decomposition and FILTERD uses a frequency decomposition providing integer values of the decimation factor. Figure 4 shows with dot lines the cutoff frequency of each frequency band that would be used for a five band uniform decomposition. In this case almost all the energy is in the first band, so the algorithm would give the cutoff frequency of this band (100 kHz) as the result. The algorithm would not detect variations of the signal bandwidth smaller than the band's width, which depends on the number of bands used. In most applications it would be enough to distinguish between the need of a low, medium, or high sample rate, so a lower number of bands can be used, providing maximum system performance.

It must be noted that analyzing a signal with this type of frequency decomposition will give bandwidth estimated values that will require the use of noninteger decimation factors to adapt the sample rate. This means using a resampling technique which is more time consuming as it requires interpolating, filtering, and decimating. Therefore another algorithm has been tested (FILTERD) based on uneven frequency decomposition by choosing the cutoff frequency of each filter so that an integer decimation factor is needed to adapt the sample rate for each band.

EXPERIMENTAL RESULTS

The system was installed at the end of March 2008 and connected to the 85 GHz reflectometer system. It has been connected in parallel with the traditional data acquisition system being used with KG8B as this project is a test bed to evaluate the applicability and performance of the proposed technique. The evaluation process will take part in several phases. The first phase started with the system installation and has had the following goals: verifying the initial hypothesis about how the bandwidth of the signals from KG8B varies during a complete shot, analyzing the reflectometer signals between 1 and 5 MHz to determine if there is any valuable information in that frequency range, and testing the algorithm behavior and performance for different scenarios. Once it is clear which the best algorithms for every situation are, and their optimum parameters have been chosen, the following phase will focus on system optimization. The plan is to have the system running during the ongoing campaign

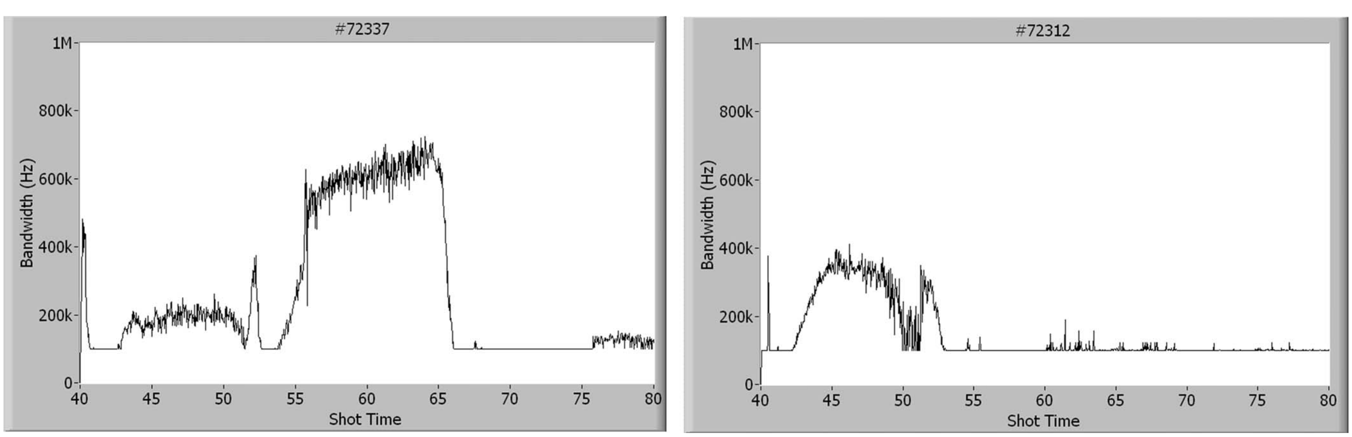

FIG. 5. Bandwidth evolution of KG8B's cosine component for the 83 GHz fixed channel for two discharges.

to demonstrate its applicability. Finally, it has been recently proposed to use this system to develop a test bed for a remote participation system that will enable users from outside JET to have live access to the data generated by this diagnostic using a web browser. To meet the first goals of the system's evaluation first phase, which are due to the limitations of the previous data acquisition systems used in KG8B, some shots have been acquired in a traditional streaming mode, and the preliminary results confirm the initial hypothesis on bandwidth evolution. Figure 5 shows the results obtained with the NOCOFE algorithm in a supervised analysis for two shots with different JET scenarios. It can be seen that the signal's bandwidth can change considerably during a shot and that it is possible to reduce the amount of data generated without losing information by reducing the sample rate. From the point of view of the validation of the system as an alternative data acquisition technique more shots have to be acquired in this mode to characterize the results for different scenarios. However, these results pose an even more important and challenging question for the near future, which establishes the link between the change in the signal's bandwidth and the variations in plasma properties. Looking at the information provided by this diagnostic from this point of view might provide additional information at the same cost.

The other main goal during this first phase has been to test and analyze the bandwidth estimation algorithms. In order to do so, an analysis application has been developed that permits to simulate the system's behavior using data previously acquired in streaming mode. The application reads the data from a file in blocks, as it would receive it during an acquisition, computes the power spectrum for each channel, and applies the four algorithms here proposed to estimate the signal's bandwidth for each block: NOCOFE, SPUR, FILTERB (uniform frequency decomposition), and FILTERD (constant D decomposition). For each block and channel, the application displays the signal's power spectrum and the bandwidth value estimated with each of the four algorithms. In addition to this it has one graph for each algorithm showing the bandwidth evolution with time for every channel up to that block. The application can run in a supervised mode allowing the user to pause the simulation at any time to analyze the results for a particular block or in an unsupervised mode where it analyzes a complete shot and saves the results in a spreadsheet file for later processing.

Preliminary results obtained with this application show that from a qualitative point of view the NOCOFE algorithm seems to be performing the better in general terms. Problems in the design of the FILTER algorithms have been detected as it has not been possible to find a threshold level that gives good and consistent results for different scenarios yet. The reason for this is the wide dynamic range of the signal, so modifications based on normalization have been included in the algorithm and are being tested. The SPUR algorithm seems to be working properly in detecting the spurious when they appear, but there are no clear criteria yet to distinguish when these spikes are noise or when they carry information. This is an important issue to treat in the near future.

TABLE I. Processing capabilities and maximum sample rate that can be achieved depending on the number of channels.

	Throughput (megasamples/s)			Maximum F_s (megasamples/s)		
No. of channels	1	2	4	1	2	4
NOCOFE	1,23	0,62	0,31	7,5	3,7	1,7
SPUR	1,23	0,62	0,31	7,5	3,7	1,7
FILTERD 5 Bands Order 4	3,03	1,43	0,71	9,3	4,3	1,8

Regarding the data reduction that can be obtained with this technique the preliminary results confirm that it has a great potential. It is difficult to give a single general figure as the reduction depends directly on the bandwidth evolution and this, in turn, depends on the type of shot. As an example, this technique would reduce the amount of data generated for shots shown in Fig. 5 by 56% and 81%, respectively.

Table I shows the impact on the system's performance that has each algorithm in terms of the maximum processing capability for each of them and the maximum sample rate that can be achieved for a 40 s acquisition depending on the number of channels used. It must be noted that the algorithm implementation has not been fully optimized in this phase as this will be done in phase two.

ACKNOWLEDGMENTS

This work, supported by the European Communities under the contract of Association between EURATOM/CIEMAT, was carried out within the framework of the European Fusion Development Agreement. Financial support was also received from the Spanish Ministry of Science and Education under Research Project No. DPI200606624, by Grant No. JC2007-00084 of the José Castillejo Program, and by Project No. M08 0020-4135 of the Universidad Politécnica de Madrid. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

[1] M. Ruiz, E. Barrera, S. López, D. Machón, J. Vega, and E. Sánchez, Rev. Sci. Instrum. **75**, 4261 (2004).
[2] G. D. Conway, G. Vayakis, J. A. Fessey, and D. V. Bartlett, Rev. Sci. Instrum. **70**, 3921 (1999).
[3] S. Hacquin, L. Meneses, L. Cupido, N. Cruz, L. Kokonchev, R. Prentice, and C. Gowers, Rev. Sci. Instrum. **75**, 3834 (2004).
[4] G. D. Conway, G. M. D.Hogeweij, M. R. de Baar, Yu. Baranov, R. Barnsley, N. C. Hawkes, X. Litaudon, J. Mailloux, E. Righi, F. G. Rimini, and K.-D. Zastrow, Plasma Phys. Controlled Fusion **44**, 1167 (2002).
[5] M. Ruiz, J. M. López, G. de Arcas, E. Barrera, R. Melendez, and J. Vega, Fusion Eng. Des. **83**, 258 (2007).
[6] E. Barrera, M. Ruiz, S. López, D. Machón, and J. Vega, IEEE Trans. Nucl. Sci. **53**, 923 (2006).